

# Modelos matemáticos e substitución lingüística

Johannes Kabatek

Universidade de Tübinga (Alemaña)

kabatek@uni-tuebingen.de

Recibido o 29 / 02 / 2012 Aceptado o 02/06/2012

## Mathematical models and language shift

### Resumo

Este artigo consiste nunha resención crítica dunha serie de traballos recentes que procuran modelar, baseándose en cálculos matemáticos, o contacto lingüístico e posibles escenarios de desprazamento lingüístico (*language shift*), é dicir, o paso do uso dunha lingua a outra, ou de morte de linguas (*language death*). Presentaremos os modelos de Abrams/Strogatz 2003, de Mira e Paredes 2005 e de Mira/Seoane/Nieto 2011, de Minett/Wang 2007 e de Castelló (2010) e Castelló et alii (2012), con referencia a diferentes situacións de contacto (incluíndo referencias ao contacto galego-castelán). Os modelos e os algoritmos dos que parten son presentados e criticados, diferenciando os graos de sofisticación e adecuación dos diferentes traballos. Identificáranse varios problemas sobre todo nos traballos de Abrams e Strogatz e de Mira e Paredes: problemas de datos, problemas de saltos argumentativos e problemas de simplificacións. Aínda así, e sobre todo con referencia aos modelos máis sofisticados, propónse continuar cunha "espiral heurística" entre modeladores e sociolingüistas para chegar a modelos máis adecuados e a unha comprensión mellor dos mecanismos que actúan no contacto lingüístico.

### Palabras chave

Sociolingüística, Epistemoloxía, Contacto lingüístico, Modelos matemáticos de contacto, Desprazamento lingüístico, Contacto galego-castelán

### Abstract

This paper presents a critical review of several recent studies that aim to model language contact, language shift and language death mathematically. We will present the models proposed by Abrams & Strogatz 2003, Mira & Paredes 2005, Mira, Seoane & Nieto 2011, Minett & Wang 2007 and Castelló (2010), as well as Castelló et al. (2012). These papers refer to different contact situations, including Galician-Spanish language contact. The different models and corresponding algorithms will be described and critiqued, distinguishing several degrees of sophistication and adequacy. The study identifies a series of shortcomings, particularly in Abrams & Strogatz' and Mira & Paredes' papers: data problems, argumentative and logical errors, and also inadequate simplifications. Nevertheless, it will be argued that the more sophisticated models in particular should continue to be developed, and a "heuristic spiral" between model developers and sociolinguists ought to be pursued, in order to achieve better models and greater insight into the mechanisms involved in language change and language shift.

### Keywords

Sociolinguistics, Epistemology, Language contact, Mathematical models of language contact, Language shift, Galician-Spanish language contact

### Sumario

1. Introducción 2. Modelos matemáticos para a descrición da dinámica en situacións de contacto 2.1. O modelo de Abrams/Strogatz 2003: o prestixio como factor determinante 2.2. A expansión do modelo de Abrams/Strogatz de Mira e Paredes 2005 e de Mira/Seoane/Nieto 2011: a relevancia da distancia interlingüística 2.3. O modelo de Minett/Wang (2007): incluíndo os bilingües 2.4. O factor da volatilidade: o modelo de Castelló (2010) e Castelló et alii (2012) 3. Modelos matemáticos e realidade sociolingüística 3.1. O problema do funcionamento dos modelos: falsas reducións e falsos supostos 3.2. Novas ideas, novas discusións 3.3. A base dos datos 4. Conclusións.

### Contents

1. Introduction 2. Mathematical models for the description of language dynamics in contact situations. 2.1. The model of Abrams & Strogatz 2003: prestige as a determining factor. 2.2. Expansion of the Abrams & Strogatz model in Mira & Paredes 2005 and Mira, Seoane & Nieto 2011: the relevance of interlinguistic distance. 2.3. Minett & Wang's 2007 model: including bilinguals. 2.4. Volatility as a factor: three models in Castelló 2010 and Castelló et al. 2012. 3. Mathematical models and sociolinguistic reality. 3.1. The problem of the function of models: false reductions and false presuppositions. 3.2. New ideas, new discussions. 3.3. The data problem. 4. Conclusions.

Este traballo nace a partir dunha colaboración con Lucía Loureiro Porto que comezou por un workshop organizado conxuntamente no marco da conferencia anual da SLE en Vilnius en setembro de 2010, con resultados publicados nun número especial do *International Journal of the Sociology of Language* (véxase Kabatek/Loureiro Porto no prelo). O meu sincero agradecemento vai tanto a ela coma aos participantes dese coloquio e aos retrucantes dos artigos do número especial. Así mesmo, agradezo os comentarios dos participantes do coloquio *Langue et linguistique catalanes* celebrado en París (Nanterre) en novembro de 2010 (onde presentamos unha plenaria sobre o tema *¿Són més leials els volàtils o els que no s'adapten? Efectes paradoxals de les nostres decisions lingüístiques*), aos meus colegas do *Tübingen Center for the Study of Language*, no que as cuestións aquí tratadas son consideradas centrais, a Juan Manuel Bello Rivas polos seus comentarios, de grande utilidade sobre todo para o aptdo. 3. e, finalmente, a dous revisores anónimos deste traballo. A responsabilidade de todos os erros e das limitacións destas breves reflexións é unicamente miña.

## 1. INTRODUCCIÓN

A lingüística do século XXI está vivindo unha auténtica revolución metodolóxica: desde distintos lados, as innovacións tecnolóxicas están permitindo novos achegamentos ao obxecto de estudo. Os métodos experimentais da psicolingüística e os achados da neurolingüística permiten saber cada vez máis sobre o procesamento cognitivo da linguaxe<sup>1</sup>, e a dispoñibilidade de grandes masas de datos lingüísticos por un lado e as posibilidades de cálculos cuantitativos case ilimitados polo outro, abren novas perspectivas en case todos os campos da disciplina. Curiosamente, toda unha serie de estudos pioneiros desas novas tendencias publicados ó longo dos últimos anos non foron realizados por lingüistas senón que naceron no ámbito da medicina e da neuroloxía, da física estatística, da bioinformática, das ciencias da cognición ou das matemáticas, con resultados publicados en revistas especializadas en física ou revistas de ciencias naturais como *Nature* ou *Science*.

¿Será que por fin chegaron os investigadores de ciencias a ensinar aos lingüistas a traballar de maneira seria e obxectiva? ¿Terán que confesar os lingüistas que en pouco tempo a ciencia “seria”, apropiándose do tema da lingua, confirmou que o eterno complexo dos “de letras” fronte aos “de ciencias” era máis que xustificable? Pensamos máis ben o contrario: o que terá que facer a lingüística actual non é quedar impresionada cegamente por datos e cálculos sen antes valorar criticamente as novas tendencias ao mesmo tempo que convirá integralas e facer delas ferramentas que nos permitan avanzar no coñecemento das cousas.

Tamén no ámbito da sociolingüística contamos, desde hai algúns anos, con varios estudos en diferentes campos que procuran explicar a dinámica do cambio lingüístico, a evolución histórica das linguas ou o comportamento de individuos nas redes sociais. Como noutros campos, tamén aquí parece difícil a colaboración entre os estudosos de diferentes áreas a pesar do obxecto de investigación común: por unha banda, os matemáticos ou físicos que elaboran modelos descoñecen xeralmente a complexidade das situacións sociolingüísticas, e pola outra banda, os sociolingüistas parecen ter dificultades coa limitación da complexidade aos parámetros reducidos necesarios para o cálculo cuantitativo. Pensamos, porén, e tal como o postularon Croft e Blythe, que a colaboración é útil:

there are benefits to collaboration between empirical researchers in the humanities, who generally have not been trained in mathematical analysis of quantitative data, and modelers from physics or other strongly mathematical fields, whose models are not always based on empirical data from the domains being modeled (Blythe/Croft 2010: 19).

Sería, porén, inxenuo agardar marabillas e milagres desa colaboración; a existencia de novos métodos non é xa de por si un resultado, é máis ben o comezo dunha ampla gama de posibilidades novas, posibilidades que haberá que tomar en serio e que esixirán novos estudos que, en último termo, poderán levar ao que noutro lugar<sup>2</sup> chamamos unha “espiral heurística”: o coñecemento dos lingüistas permite, en colaboración con matemáticos ou físicos, a creación de modelos de cálculo da dinámica social e lingüística, que permitirán formular novas preguntas e que pola súa parte axudarán, en retorno á realidade social, a afinar o modelo, volvendo a partir da versión refinada á realidade social, e así sucesivamente. No vaivén entre modeladores e lingüistas, á condición de que sexa unha colaboración aberta e construtiva, o noso coñecemento sobre a realidade aumentará paulatinamente, refinando tanto os modelos e os parámetros relevantes que entran neles como os estudos de base que se necesitan para unha modelación adecuada.

<sup>1</sup> Cf. Bornkessel-Schlesewsky/Schlesewsky (2009).

<sup>2</sup> Vid. Kabatek/Loureiro Porto (no prelo).

As liñas que seguen teñen como obxectivo presentar algunhas reflexións no sentido desa “espiral heurística”: no seguinte apartado, presentaremos catro modelos de cálculo de dinámicas lingüísticas en situacións de contacto. Trátase de modelos que procuran describir escenarios de competición entre dúas linguas, coas posibles consecuencias de desprazamento lingüístico (*language shift*), é dicir, o paso do uso dunha lingua a outra, ou ata de morte de linguas (*language death*). Na sección 3, analizaremos de maneira crítica os resultados principais deses modelos, discutindo cales son os problemas fundamentais que supoñen. Por último, ofreceremos algunhas pautas para futuras investigacións. O ámbito lingüístico ao que nos referiremos será en primeiro lugar a situación sociolingüística galega, aínda que unha grande parte das reflexións ofrecidas é tamén aplicable a outras situacións de contactos lingüísticos ou incluso de competición entre variedades dunha mesma lingua.

## 2. MODELOS MATEMÁTICOS PARA A DESCRICIÓN DA DINÁMICA EN SITUACIÓNS DE CONTACTO

### 2.1. O modelo de Abrams/Strogatz 2003: o prestixio como factor determinante

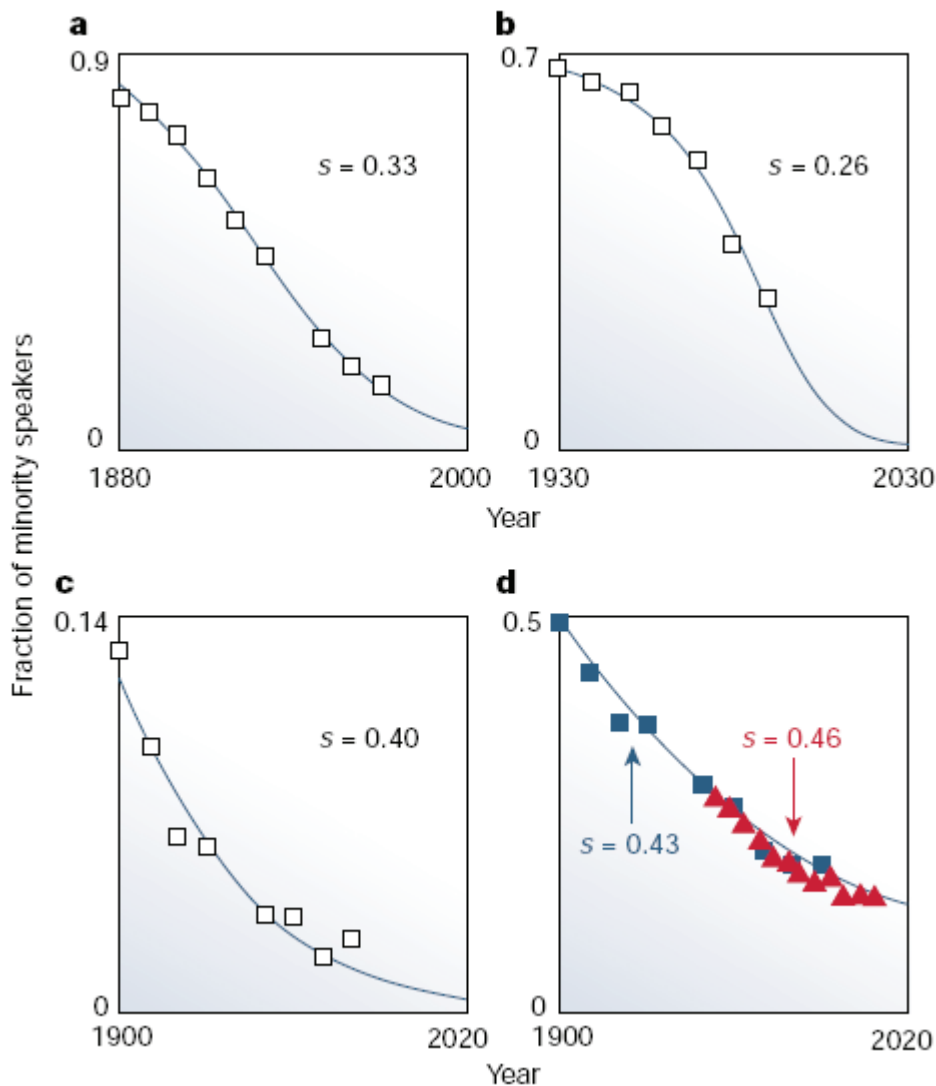
En 2003, os matemáticos Daniel Abrams e Steven Strogatz publicaron en *Nature* un traballo titulado “Modelling the dynamics of language death” no que un modelo simple de dinámica de linguas en contacto é aplicado a 42 situacións diferentes en diversos países. O resultado sorprendente deste citadísimo traballo é que unha fórmula moi simple con moi poucos parámetros é capaz de dar conta dunha grande variedade de situacións de contacto, calculando a probabilidade na que se vai superpoñer unha lingua sobre outra en situacións de contacto de dúas linguas. A fórmula coa que traballan os autores é a que se ve na fig. 1.

$$\frac{dx}{dt} = yP_{yx}(x,s) - xP_{xy}(x,s)$$

**Fig. 1.** Modelo “mínimo” de cambio lingüístico (i.e. aquí realmente de desprazamento lingüístico) segundo Abrams/Strogatz 2003

Considéranse aquí dúas linguas X e Y, que entran en competición, tendo un status ou prestixio (s) variable. O que calcula o modelo é a probabilidade coa cal os falantes cambian de Y a X. Abrams e Strogatz chaman o seu modelo “mínimo” e só teñen en conta falantes monolingües nunha comunidade considerada homoxénea, é dicir, con contactos a través de toda a comunidade, sen variación xeográfica ou doutro tipo.

No seu breve traballo, os autores só presentan algúns dos resultados, pero indican que en realidade testaron 42 situacións lingüísticas, inserindo no modelo os datos dos censos: cando dipoñibles, datos sobre a porcentaxe poboacional respectiva que fala cada unha das linguas en cuestión, e cando non, extrapolando a partir de datos sincrónicos con informacións sobre diferentes grupos de idade, ou aproveitando indicadores indirectos, coma as misas celebradas en quechua no caso de Huanuco, Perú. Algúns dos resultados vense na fig. 2:



**Fig. 2.** Dinámica de morte de linguas segundo Abrams/Strogatz 2003. A situación *a* describe a evolución (i.e., o número de falantes) do gaélico escocés en Sutherland; *b* a do quechua en Huanuco, *c* o galés en Monmouthshire e *d* o galés en todo Gales. Os datos dos cadrados en *d* derivan de datos históricos, os datos dos triángulos son extrapolacións a partir dun censo moderno.

O resultado máis destacado é que o parámetro  $s$ , é dicir, o prestixio dunha lingua, é o máis relevante para a súa evolución:

status,  $s$ , is the most relevant linguistically; it could serve as a useful measure of the threat to a given language (Abrams/Strogatz 2003: 900).

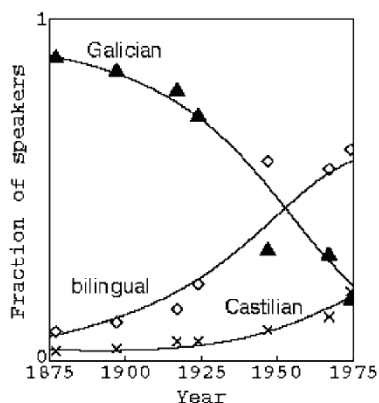
O modelo predí —e esta é a hipótese forte da que parte— que non poden existir sociedades bilingües estables, é dicir, que segundo ese modelo non existe o chamado “bilingüismo harmónico”, aínda que tamén asume que a realidade pode contradicir esa hipótese, coma no

caso do Québec, pero amósase, non obstante, que isto é debido ao feito de que non haxa moita interacción entre os dous grupos practicamente separados: o Québec sería entón un caso de dúas comunidades máis ou menos monolingües, convivindo no mesmo territorio pero sen a mestura necesaria para que haxa realmente competición. Tamén se chega á conclusión de que a desprazamento lingüístico (*language shift*) pode impedirse ou frearse con medidas políticas: se na fórmula se introduce un control sobre o parámetro  $s$ , de prestixio (é dicir que se mediante medidas de planificación de estatus,  $s$  sobe), pode chegarse a estabilizar a situación.

Deixemos de momento aberta a cuestión da validez e o valor explicativo (ou incluso predictivo) deste modelo para a descrición adecuada de situacións de substitución lingüística e pase-mos á presentación de tres modelos máis antes de comentalos no seu conxunto.

## 2.2. A expansión do modelo de Abrams/Strogatz de Mira e Paredes 2005 e de Mira/Seoane/Nieto 2011: a relevancia da distancia interlingüística

En 2005, dous físicos da Universidade de Santiago, Jorge Mira e Ángel Paredes, presentaron unha versión revisada do modelo de Abrams e Strogatz baseada explicitamente no caso do contacto castelán-galego. A hipótese fundamental dese traballo é que, en contra de Abrams e Strogatz, o bilingüismo estable si é posible, pero o factor do que depende isto é a similitude entre as linguas. Partindo da observación da evolución da situación lingüística galega dos 150 anos antes do final do franquismo (é dicir, antes do comezo de intervencións conscientes en prol do galego, ver fig. 3), os autores observan un crecemento continuo de bilingües que acompaña o declive do galego.



**Fig. 3.** Dinámica da evolución do galego, segundo Mira/Paredes 2005 baseado, segundo afirman os autores, en datos do Seminario de Sociolingüística da RAG<sup>3</sup>.

Para modelar os datos da fig. 3 co modelo de Abrams/Strogatz, esíxese a introdución dun factor non previsto no modelo orixinal: os bilingües. De feito, Mira e Paredes sospeitan que a falta de bilingües no modelo de Abrams e Strogatz é debida á escolla das situacións descritas polo modelo: situacións de contacto entre linguas tipoloxicamente distantes e sen posible intercomprensión. Introducendo os bilingües no modelo, introducen, ademais, outro parámetro que é a distancia entre as linguas  $k$ , que pode variar entre 0 (distancia moi alta) e 1 (distancia cero ou identidade das dúas linguas en competición). A posibilidade dos falantes de escollerem a opción do bilingüismo e non a do cambio de lingua aumenta, segundo este modelo, conforme aumenta a similitude entre as linguas en cuestión. Os autores resumen este resultado así:

<sup>3</sup> Sobre o problema deses datos, véxase a nota 22.

we have shown that, in a model of competing languages, bilingualism is possible. Similarity appears to be the key factor that enables the stability of bilingualism. (Mira/Paredes 2005, 1033)

En 2011, unha versión ampliada deste traballo foi publicado por J. Mira en colaboración con L. F. Seoane e J. J. Nieto na revista en liña alemá *New Journal of Physics* ([www.njp.org](http://www.njp.org))<sup>4</sup>. Nesta ocasión, presentouse un traballo máis elaborado, con información máis detallada e que incluía unha serie de suxestións interesantes, como a de ter en conta alteracións da situación inicial ao longo dun proceso (por exemplo, variacións do índice de nacemento ou cambios de prestixio) ou a de diferenciar espazos ou modelar axentes en vez de sistemas. Aínda así, a idea fundamental xa presentada no estudo de 2005 é mantida: o factor determinante para a posibilidade de mantemento dunha lingua en situacións de contacto é o parámetro da similitude interlingüística ("interlinguistic similarity", p. 3).

### 2.3. O modelo de Minett/Wang (2007): incluíndo os bilingües

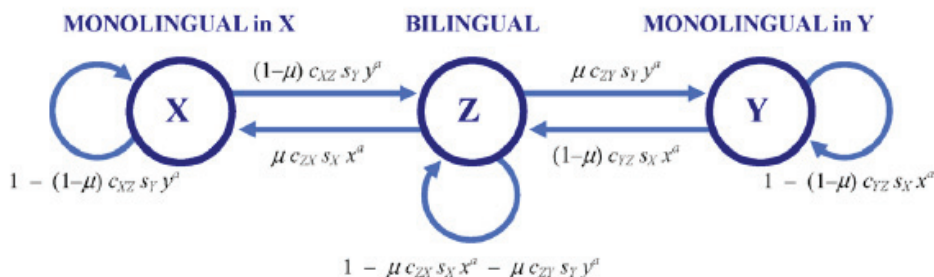
Xa catro anos antes do traballo de Mira/Seoane/Nieto, dous investigadores da Chinese University de Hong Kong publicaron, na revista *Lingua*, un traballo relativamente exhaustivo que non só inclúe, fronte a Abrams e Strogatz, o papel dos bilingües, senón que ademais presenta un modelo baseado na actividade dos axentes e non só nas características globais dos sistemas. Este traballo, aínda que inicialmente atopemos algúns tópicos sobre as linguas do mundo que son constantes nos estudos aquí presentados<sup>5</sup>, difire considerablemente dos anteriores xa que emerxe da colaboración entre un lingüista e un matemático, en concreto, do matemático James Minett co lingüista William Wang. Isto leva a que a información sobre principios sociolingüísticos sexa aquí moito máis precisa que en Abrams e Strogatz ou nos artigos dos físicos galegos. Ademais, é un traballo máis cauteloso á hora de definir os factores cos que operan os modelos; fálase, entre outros, do factor "attractiveness", procurando non introducir interpretacións demasiado atrevidas. Aínda así, mesmo se o punto de partida é precisamente a crítica das *simplificacións* do modelo de Abrams e Strogatz (Minett/Wang 2007, 21), atopamos tamén aquí a necesidade de partir dunha serie de reducións da realidade<sup>6</sup>.

O modelo de Minett e Wang, un traballo puramente teórico sen aplicación empírica, amplía o modelo de Abrams e Strogatz en dous sentidos, integrando, por unha banda, o factor dos bilingües, e, pola outra, o da estrutura social. Argumentando que o paso dun falante dunha lingua A a unha lingua B sen estadios intermedios é un fenómeno bastante pouco probable, e que na realidade haberá, nese proceso de desprazamento lingüístico, algún momento de competencia (aínda que sexa limitada) nas dúas linguas de contacto, os autores parten de dous tipos de desprazamento: o desprazamento vertical (entre xeracións) no que os factores determinantes son a mortalidade, a taxa de nacemento, e a probabilidade de transmisión dunha das dúas linguas de contacto ou de ambas, e o modelo horizontal, de contacto entre falantes adultos. Os dous submodelos (vertical e horizontal) intégranse nun modelo sinóptico (véxase fig. 4).

<sup>4</sup> Este artigo tivo un amplo eco na prensa galega, vid. entre outras *La Voz de Galicia* do 26-I-11, onde apareceu un artigo co título "Las matemáticas demuestran que el gallego tiene futuro". Tamén existe un vídeo en Youtube que resume o traballo: <http://www.youtube.com/watch?v=oOotTKZY-Is>

<sup>5</sup> Poderíase facer unha análise crítica do discurso dos artigos presentados aquí: todos comezan falando da extinción das linguas do mundo citando xeralmente as mesmas fontes, e todos espertan unha certa expectativa no lector cando falan da posibilidade de analizar obxectivamente as situacións de linguas en perigo ou das medidas terapéuticas para a salvación das mesmas.

<sup>6</sup> p. ex. "Throughout the current section we will assume that the agents are all neighbours of each other, modelling a population having a fully connected social network." (Minett/Wang 2007: 34).



**Fig. 4.** Modelo unificado de contacto de Minett/Wang 2007, incluíndo bilingües. Cálculase a probabilidade de mantemento de  $x$ , de paso de  $x$  a  $z$ , de mantemento de  $z$  e do paso de  $z$  a  $x$  e  $a$   $y$ , así como de mantemento de  $y$  ou do paso de  $y$  a  $z$ . Exclúese o paso directo dunha lingua a outra sen pasar por  $z$ , e intégrase tanto a transmisión vertical como a horizontal

Minett e Wang mostran en varias visualizacións como a variabilidade dos factores “atractividade de lingua” e “cantidade de monolingües” en combinación co factor “mortalidade”/contacto horizontal (respectivamente: dinamicidade do sistema) xera un sistema vectorial que pode levar a diferentes resultados. Con referencia tanto a traballos de Fishman e á escala GIDS de continuidade de uso dunha lingua nunha comunidade (Fishman 1991 e 2001) coma a David Crystal (2000) e os seus seis mecanismos fundamentais para salvar unha lingua, Minett e Wang indican que tanto as intervencións no *status planning* que chegan a aumentar o prestixio dunha lingua como o aumento de falantes monolingües mediante cambios no mundo educativo modifican o *output* do sistema e axudan a manter unha lingua. Fronte a Abrams e Strogatz, o modelo de Minett e Wang parece demostrar, pois, que o mantemento dunha lingua minoritaria é posible se o factor fundamental do seu atractivo está garantido.

Nun segundo modelo cambian a perspectiva desde o sistema aos axentes, é dicir os individuos que compoñen o sistema e que o forman<sup>7</sup>. En vez de supoñer que todos os membros dun sistema interactúan da mesma maneira, o modelo de axentes permite a inclusión do factor da estrutura social xa que non todos os axentes teñen contacto con todos: suponse, en cambio, que hai individuos máis conectados na rede social e outros con menos conexións. Como non traballan con datos empíricos, os autores non queren e non poden demostrar a maior validez do modelo de axentes, pero sinalan que este claramente está máis preto da realidade xa que permite ter en conta máis factores.

Conclúen o seu traballo suxerindo unha serie de melloras que aínda se poderían integrar no modelo, como a diversidade xeográfica (no modelo pártese da situación pouco probable de homoxeneidade xeolingüística) ou o grao de intelixibilidade mutua, factor crucial do modelo presentado na sección anterior.

## 2.4. O factor da *volatilidade*: o modelo de Castelló (2010) e Castelló et alii (2012)

Terminaremos o noso repaso de modelos coa presentación dun derradeiro caso, desenvolvido polo laboratorio de física estatística da universidade das Illas Baleares dirixido por Maxi San Miguel, levado ao cabo sobre todo por Xavi Castelló na súa tese de doutoramento de 2010<sup>8</sup>, e presentado ao público anterior e posteriormente en varios artigos escritos en colaboración e coa participación, entre outros, da lingüista Lucía Loureiro Porto.

<sup>7</sup> “Unlike dynamical systems [...] in which the global behaviour of the entire system is determined by a single set of differential equations, the dynamics of agent-based models are described locally in terms of how individual agents interact with each other.” (Minett/Wang 2007, 34; Cf. tb. Axelrod 2006).

<sup>8</sup> Castelló 2010.

Igual que Minett e Wang, Castelló et al. parten de tres grupos, falantes monolingües dunha lingua A, falantes monolingües de B e falantes bilingües de AB. Os falantes interactúan nunha rede social (*social network*), e esa rede pode presentarse de maneira variada no modelo, con conexións múltiples entre os axentes ou con relacións máis reducidas. Pero o que é novo con respecto aos modelos anteriores presentados non é só a maneira máis sofisticada de ter en conta diferentes tipos de estrutura social e de posibilidades de interacción dos axentes, senón, ademais, a inclusión dun parámetro distinto que dá conta do factor do prestixio encuberto (*covert prestige*, cf. Trudgill 1972): sábese que o factor “prestixio” do que parten Abrams e Strogatz é altamente problemático xa que en moitas comunidades, aínda que o prestixio aberto dunha lingua sexa moi baixo, hai unha lealdade moi alta a esta lingua, sobre todo en certos ámbitos (familia etc.), o que pode permitir a súa supervivencia a longo prazo (pensemos por exemplo en casos como certos dialectos do alemán ou do italiano, de baixo prestixio aberto pero totalmente vivos a pesar diso; ou na lingua galega e a súa plurisecular supervivencia como lingua de baixo prestixio aparente). Para dar conta desa aparente contradición entre prestixio e mantemento (devida ao prestixio encuberto), o novo parámetro incluído no modelo por Castelló et al. é o da lealdade lingüística ou *volatilidade* (“volatility”), é dicir a dispoñibilidade dos individuos axentes de cambiar de lingua. A baixa volatilidade, segundo estes autores, é debida ou ben á resistencia por parte dos falantes a cambiar de lingua, ou ben, en situacións como as descritas por Mira e Paredes de parentesco alto entre as linguas, á posibilidade de comprensión mutua<sup>9</sup>.

Un dos resultados máis chamativos do modelo<sup>10</sup> é o efecto do parámetro volatilidade: a baixa volatilidade pode contribuír a retardar o efecto de extinción dunha lingua xa que hai menos contacto. Pero curiosamente (e os autores falan dun efecto paradoxal), é sobre todo a alta volatilidade a que conserva, segundo o modelo, as linguas<sup>11</sup>. Se os falantes mostran unha alta dispoñibilidade a cambiar de lingua, contribúen a que as dúas linguas de contacto sigan sendo faladas e a extinción da máis débil se pospoña.

Sen poder entrar en todos os detalles dos traballos deste grupo (como a cuestión da estrutura social e a cuestión de análises cualitativas e cuantitativas), podemos resumir que a modelización sobre a base de diferentes parámetros relacionados co comportamento individual pode facer ver sumas sociais dese comportamento que parecen sorprendentes: en contra da intuición bastante xeral entre falantes de comunidades con presenza dunha lingua minoritaria de que o mellor que se pode facer para fomentar a lingua minoritaria é falala en todas as circunstancias posibles da vida, o modelo parece indicar outra cousa: ¿será realmente que un comportamento decididamente “volátil”, con cambios de lingua frecuentes e alta dispoñibilidade de adaptarse aos demais contribuír, en contra do que se adoita pensar, ao mantemento da lingua minoritaria?

Discutiremos na seguinte sección os alcances e límites deses modelos, non sen sinalar que unicamente escollemos aquí uns cantos exemplos e que hai outros modelos, por exemplo aqueles baseados en diferentes ramas da teoría dos xogos, que tamén se ocuparon do tema da modelización do contacto lingüístico. Pero a finalidade do presente traballo non é a de ofrecer unha visión exhaustiva de todos os traballos existentes nesa vertente senón de preguntarnos polo proveito que a sociolingüística pode sacar deles.

<sup>9</sup> Como dicía unha das informantes de Kabatek 2000: “A: N/ non, non, .. ¿non é por radicalismo! é que como o galego o entende a xente, .. entón por que vas estar cambiando de lingua ¿non?”

<sup>10</sup> Atópase unha aplicación informática en internet que permite traballar co modelo de bilingües de Castelló fronte ao modelo de Abrams/Strogatz, con posibilidade de introducir diferentes valores para os diferentes parámetros, véxase [http://ifisc.uib-csic.es/research/complex/APPLET\\_LANGDYN.html](http://ifisc.uib-csic.es/research/complex/APPLET_LANGDYN.html)

<sup>11</sup> “high volatility (small a) is shown to generally favor coexistence” (Castelló 2010, 162).



### 3. MODELOS MATEMÁTICOS E REALIDADE SOCIOLINGÜÍSTICA

Os traballos que presentamos na sección anterior só supoñen unha pequena mostra que podería ser completada por outros: case podemos falar dunha moda de modelar situacións de linguas en contacto nos últimos anos, moda que forma parte dunha tendencia máis ampla a modelar sistemas dinámicos do mundo natural, social e cultural, con toda unha serie de traballos sobre cuestións de linguas naturais<sup>12</sup>. Neste apartado, preguntáronos polos alcances e límites dos modelos presentados. A perspectiva haberá de se converter, pois, nunha perspectiva máis crítica.

Comecemos polo lado dos alcances: ¿que é o que sabemos sobre situacións de contacto lingüístico que non sabiamos antes? ¿Cal é a achega dos estudos presentados ao coñecemento sobre o comportamento dos falantes e sobre os sistemas sociais en xeral? A resposta deberá ter en conta o diferente grao de complexidade dos modelos e, por conseguinte, o diferente grao de adecuación á realidade. E deberá ter en conta, así mesmo, o grao de cautela á hora de presentar os resultados.

O primeiro resultado positivo de todos os traballos presentados dáse pola súa configuración: todos se basean en algoritmos relativamente simples que funcionan basicamente aplicando reiteradamente unha función con algunhas variables. As variables son moi poucas, a aplicación do algoritmo, a reiteración, en cambio, é interminable. É dicir que se pode calcular o efecto colectivo masivo, en principio case sen límite cuantitativo, de mecanismos moi simples. Na sociolingüística contamos, por unha banda, coa tradición microsociolingüística, na que múltiples factores se describen minuciosamente e, pola outra banda, coa tradición macrosociolingüística, na que poucos factores son considerados relevantes para a análise cuantitativa. Os modelos descritos neste traballos inscribíense claramente na última liña de tradición, coa que comparten a hipótese de que a aparente complexidade dos sistemas sociais é reducible a poucos factores relevantes<sup>13</sup>. Alén da “técnica” do cálculo, os modelos esixen, pois, a máxima redución de factores e unha profunda reflexión sobre cales son os factores fundamentais que rexen o noso comportamento. Esa reflexión —que pode tamén desembocar no resultado de que unha tal simplificación é imposible— é altamente pertinente e corresponde a unha das esixencias máis fundamentais do traballo científico; aínda así, a redución dunha enorme complexidade a poucos factores relevantes sempre é tamén perigosa e provocará crítica<sup>14</sup>.

O segundo resultado fundamental xa non é común a todos os modelos, senón que se limita a aqueles que traballan co resultado colectivo do comportamento dos axentes: as linguas e a dinámica do contacto non son realidades abstractas independentes, senón resultado da actividade dos falantes. O cambio de perspectiva do sistema aos axentes corresponde ao postulado de non mirar cara a unha “man invisible”<sup>15</sup> cando se fala da dinámica lingüística, senón de mirar á actividade visible dos falantes, únicos axentes reais nos procesos de dinámica social.

O terceiro resultado positivo pode parecer negativo a primeira vista, pero convértese en positivo tendo en conta a idea da “espiral heurística” mencionada máis arriba: a necesidade de simplificación dos factores será criticada se sobrepasa os límites do reducible ou se atribúe ex-

<sup>12</sup> Unha lista máis exhaustiva de traballos publicados por físicos sobre linguas naturais atópase baixo <http://www.pks.mpg.de/~edugalt/physicist-language/>

<sup>13</sup> Como din Castelló et al. (*no prelo*): “our current understanding of complex systems shows broad evidence that complex behaviour can arise from very simple interaction rules, and that collective behaviour shows universal features.”

<sup>14</sup> Pénlese por exemplo na idea freudiana de reducir gran parte do comportamento humano á libido: unha idea que foi acollida como fascinante e acertada, por algúns, e rexeitada polo seu enorme reduccionismo, por outros. En sociolingüística, podemos distinguir dous tipos de reducións: por unha banda, as reducións de factores, cando por exemplo se di que o que determina a escolla dun elemento ou dunha lingua é unicamente o prestixio desa forma ou lingua, e, por outra banda, a redución a elementos considerados representativos, cando por exemplo se considera que un fenómeno, coma a pronuncia do /r/ nos famosos estudos de Labov sobre Nova York, é un indicador para a realidade social. Ambas as reducións son perigosas xa que presupoñen un coñecemento profundo, dificilmente sistematizable, das situacións en cuestión: dalgún xeito, os investigadores, para chegaren a tales reducións, necesitan xa saber de antemán que elas son as relevantes.

<sup>15</sup> Esa imaxe introducida na teoría do cambio lingüístico por Keller (1994) foi criticada entre outros por Kabatek 2005.

plicacións problemáticas a certos factores. As reaccións críticas levarán a propostas de mellora, e producirán reaxustes ou modificacións dos modelos que acrecentarán a validez dos mesmos —á condición de que haxa consenso sobre a principal posibilidade de representar a dinámica lingüística mediante modelos—.

Con estas últimas observacións xa estamos pasando aos límites dos modelos, e comezaremos por dúas observacións xerais: en primeiro lugar, os lingüistas non deberían deixarse intimidar polo aparato fórmulas dos modelos presentados; deben intentar comprender de que maneira están feitos realmente e que é o que representan, en vez de pensar que son caixas negras nas que un introduce un problema non resolto por un lado e recibe a solución polo outro. Non debemos, en definitiva, deixar a comprensión dos modelos en mans dos modeladores fiándonos dos resultados e pensar que os modeladores xa saberán o que fan por seren eles os expertos nos cálculos: os modelos non farán máis ca aplicar algoritmos aos datos que se introducen neles; poden axudar a extrapolar en cantidades grandes algunhas tendencias, pero non fan milagres. Só unha vez que saibamos realmente como funcionan, podemos chegar a unha crítica construtiva que permite engadir factores que acheguen máis o modelo á realidade.

En segundo lugar, paréceme que é difícil, nalgúns dos traballos presentados, chegar aos resultados propostos coa base dos datos lingüísticos e sociolóxicos empregados, e non habería que deixarse levar polos cálculos e as cifras pensando que terán unha obxectividade intrínseca: non habería que caer no eterno complexo da “xente de letras” fronte á suposta “verdadeira cientificidade” das ciencias chamadas exactas<sup>16</sup>. En parte, a crítica debe, ao noso ver, indicar os problemas de datos dos que padecen os traballos presentados e propor, na medida do posible, melloras neste sentido. Non é que os físicos e os matemáticos souberan de repente máis de lingua que os sociolingüistas. De feito, nalgúns dos traballos presentados, é precisamente a falta de coñecemento lingüístico a que leva a resultados problemáticos, e o remedio consistiría nunha verdadeira colaboración entre os modeladores e os lingüistas (ou na formación dunha nova xeración de lingüistas expertos nos dous campos)<sup>17</sup>.

### 3.1. O problema do funcionamento dos modelos: falsas reducións e falsos supostos

Un dos discursos actuais no ámbito da ciencia da linguaxe apunta a que chegou por fin a hora de aplicar os estándares das outras disciplinas científicas tamén na lingüística. Enténdese, entre liñas, que a lingüística é unha ciencia atrasada, que traballa con intuicións, con datos pouco obxectivos e con xuízos non obxectivables. Dun xeito semellante ao que aconteceu no final do século XIX coa busca da obxectividade científica por parte dos Neogramáticos, hoxe en día búscase a obxectividade na neurolingüística ou na lingüística cuantitativa. Xa dicíamos que a mera cuantificación non dá de por si máis obxectividade; e hai que engadir que a adopción dos paradigmas das ciencias naturais en lingüística é problemática xa que, contrariamente ao que pasa nas ciencias naturais, nas ciencias humanas contamos cun coñecemento intuitivo previo dos obxectos que non debería verse como problemático ou como motivo de despiste do investigador que non é capaz de saír do cárcere do obxecto que está investigando: todo o contrario, ese coñecemento intuitivo permite unha visión sobre as cousas imposible nas ciencias naturais, e debe considerarse como unha vantaxe enorme na que se pode basear todo estudo científico, iso si, non sen “obxectar” a intuición subxectiva. Hai que atopar, en cada estudo, unha metodoloxía adecuada para pasar da intuición subxectiva a unha visión intersubxectivamente comprobable dos feitos. E hai que ser consciente da parte correspondente á intuición en cada

<sup>16</sup> As cales, ben mirado, tampouco son tan exactas; véxase, por exemplo, Katsumori 2010.

<sup>17</sup> Un exemplo é a proposta do *Tübingen Center for the Study of Language*, onde esta combinación é precisamente parte do programa de formación de futuros lingüistas.

estudo. De feito, é sabido que moitas veces os traballos nas humanidades que negan o papel da intuición tamén están baseados nela, e un dos maiores erros do behaviorismo lingüístico, por poñer un exemplo da historia da lingüística, foi o de non darse de conta de que, alén de toda observación pura do comportamento sen entrar na mente dos falantes, sempre houbo momentos de intuición e de mentalismo. Algo semellante acontece no caso dos modelos matemáticos: nos traballos presentados, atopamos, por unha banda, reducións ás veces problemáticas da realidade lingüística, e atopamos, ademais, saltos argumentativos nos que a intuición está presente sen que isto se diga explicitamente. No traballo de Abrams/Strogatz (2003), a redución máis rechamante consiste na idea da existencia dunha sociedade de dous grupos completamente monolingües. O modelo calcula a transición de individuos dun grupo ao outro sen ningunha fase intermedia. Isto é evidentemente algo inexistente na realidade. Así mesmo, en varios dos estudos trabállase con comunidades supostamente homoxéneas, ou asúmese que todos os falantes están en contacto con todos do mesmo xeito, ou suponse que o status relativo de dúas linguas de contacto é estable ao longo do tempo, etc. Aquí, os lingüistas deberían intervir e criticar que non se pode partir de modelos que “supoñen” algo que non corresponde á realidade xa que un invento artificial dun modelo tampouco terá o potencial de representar a realidade. Hai reducións que son posibles, pero haberá que ter moito coidado á hora de analizar resultados que se deberon ás reducións indebidas.

O problema dos saltos argumentativos dáse cando de repente unha observación baseada en datos obxectivos é tomada como base para afirmar outra cousa distinta, facéndolle crer ao lector que o resultado deriva inmediatamente da observación dos datos sen intervención doutro tipo. De feito, os datos só son os datos e non din máis do que din, e o cálculo pode corresponder á súa evolución, pero non “explica” nin “obxectiviza” os fenómenos reais. Así, por exemplo, a única evidencia do traballo de Abrams e Strogatz (2003) é que crea unha fórmula que chega a visualizar matematicamente unha liña que corresponde á liña de evolución de dous grupos de falantes en competencia. Pero esa fórmula non “explica” esa liña nin tampouco ofrece o parámetro responsable de tal evolución. Se no traballo de Abrams e Strogatz se fala do prestixio como factor determinante, isto é porque os autores teñen unha intuición (e saben desde a literatura relevante) da importancia do prestixio na competición das linguas. Pero non é que os datos digan “obxectivamente” que o responsable da evolución sexa o prestixio. Só din que no contacto de dúas linguas, o proceso de desprazamento dunha das dúas vai seguindo ritmos diferentes segundo a situación concreta e que a diferenza de ritmo se pode calcular matematicamente. Nada máis.

De xeito aínda máis problemático, no traballo de Mira e Paredes de 2005, a idea de que o factor determinante da emerxencia dun grupo de bilingües é a proximidade das linguas en contacto, *non* deriva en realidade do modelo: o único que fai o modelo é calcular que un tal grupo emerge, sen saber por que. No inicio do artigo, os autores aínda din que *suxiren* (“suggest”) que a similitude entre as linguas —nese caso o castelán e o galego— pode ser un factor determinante para a emerxencia dun grupo de bilingües, modalizando que dela *podería* (“may”) depender a estabilidade do sistema. Isto evidentemente é pura intuición e non corresponde a ningunha obxectividade, pero os autores non o din. Ao final, chegando ás conclusións, din que demostraron a posibilidade dun bilingüismo estable e que o que aparece como factor determinante do bilingüismo é a similitude entre as linguas. Todo isto non está nos datos: os datos unicamente visualizan que xurdiu un grupo de bilingües e que o grupo dos monolingües galegos diminuíu. Nada máis. Todo o resto é interpretación. E isto tampouco cambia no traballo máis actual de Mira, Seoane e Nieto publicado en 2011: o parámetro  $k$  no modelo representa a probabilidade de que a desaparición dun falante monolingüe non teña como resultado o salto á outra lingua senón a adopción dunha práctica bilingüe, cousa que parece convincente e de acordo con

observacións tradicionais da socioloxía da linguaxe. Aínda así, nada xustifica o paso seguinte, cando os autores din: “We identify interlinguistic similarity with this parameter  $k$ ” (Mira/Seoane/Nieto 2011, 3). Non explicitan de onde quitan esta información. “Identifican” este factor. ¿Pero como? ¿Hai algunha base teórica ou empírica para esta observación? Se eu afirmo, en cambio, que ese factor en realidade vén determinado polo “desexo de ser bilingüe” ou se identifico incluso outros indicios como “falta de carácter” ou “consumo de viño branco” como factores determinantes, pode parecer máis ou menos razoable ou absurdo, pero está baseado no invento espontáneo igual que o da semellanza. Ademais, ¿como se pode saber que a similitude de linguas é o factor determinante se unicamente se presenta un estudo dunha única situación? En realidade, o que se fai aquí é simplemente engadir un par de cálculos máis a un modelo de dúas liñas de evolución para que haxa unha terceira liña emerxente. Despois búscase na experiencia (os autores coñecen de cerca a situación galega) o que *podería* ser o factor determinante. Agora ben, antes de dar por comprobado que o factor “inventado” é o relevante, habería que poñer sobre a mesa diferentes factores posibles e ver como estes se poden avaliar, dándolle prioridade a un deles e non a outro. Pero isto non se resolve con cálculos senón desde a empatía, o bo coñecemento e a análise pormenorizada por parte do lingüista.

### 3.2. Novas ideas, novas discusións

Pero non todo é invento, non todo é circular: os modelos tamén poden facer ver consecuencias colectivas do comportamento individual que desde unha perspectiva individual non se ven facilmente. Os individuos, por estarmos dentro de nós mesmos, temos sempre dificultades á hora de extrapolar; procuramos sempre vernos por fóra, pero non podemos: estamos no cárcere da nosa propia mente. Son coñecidos os exemplos dos resultados colectivos involuntarios do comportamento individual, como o atasco de tráfico ou a caída da bolsa: ningún dos axentes adoita sentirse culpable do resultado e incluso adoita botarlles a culpa aos demais, pero cada acto mínimo de frear contribúe ao atasco tal como cada venda de accións contribúe á caída da bolsa. O ámbito onde o comportamento individual nos engana e non vemos a que resultado colectivo estamos posiblemente contribuíndo paréceme o de máis potencial dos modelos. Neste sentido, é moi provocador o parámetro da “volatilidade” identificado polo grupo de Castelló et al. Non se trata aquí dun parámetro inventado como o da proximidade das linguas ou o do prestixio: trátase, en cambio, dun nome dado a algo que realmente está no modelo, é dicir, a dispoñibilidade dun axente a adaptarse ou non ao comportamento dos axentes cos que entra en contacto.

O feito de que o modelo demostra que a alta volatilidade estabiliza o sistema ou retarda a desaparición da lingua B no contacto de dúas linguas parece contradicir a intuición da maioría dos falantes de linguas minoritarias, e vai en contra, ademais, do comportamento práctico de moitos falantes “comprometidos” cunha lingua minoritaria que parten do suposto de que un comportamento canto máis monolingüe posible na lingua B, máis vai contribuír ao mantemento da mesma. Así, por exemplo, a idea da maioría dos falantes galegos que queren fomentar o uso do galego é que falar sempre galego loxicamente é o que máis vai axudar a aumentar o uso da lingua. ¿E agora de repente chega un modelo matemático que nos di que todo este comportamento foi errado e que tería sido mellor cambiar sempre ao castelán cando alguén fala castelán, como parecen mostrar, por exemplo, as cifras presentadas na fig. 5?

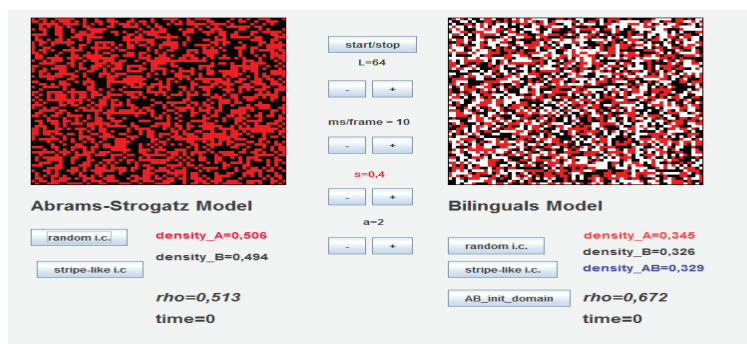
A resposta a esta observación paradoxal é dobre; dunha banda, hai que dicir que o modelo non calcula o factor volatilidade como factor asimétrico e que na realidade é pouco probable atopar situacións onde a volatilidade sexa simétrica. É dicir, por exemplo, que me parece moito máis probable que un falante de galego en Galicia cambie ao castelán en contacto cun castelanfalante ca o contrario (aínda que o contrario tamén acontece a miúdo). Mentres a volatilidade sexa asimétrica e tendencialmente sexan os falantes da lingua B os que se se adapten á lingua A pero non

ao revés, resulta claro que o efecto da volatilidade vai reducir o uso de B, e de aí vén a estratexia común entre os falantes de linguas consideradas minorizadas de reaccionar en contra da perda desas linguas cun comportamento pouco volátil e fiel á lingua sentida como ameazada. Habería, pois, que introducir no modelo a posibilidade de programar de maneira asimétrica ese factor.

---

**APPLET: Abrams-Strogatz Model VS Bilinguals model**

---



**Fig. 5.** Aplicación en liña para a comparación do funcionamento do modelo de Abrams/Strogatz 2003 e o dos bilingües de Castelló 2010 (cf. nota 11). O factor  $s$  (prestixio) neutro é 5; o factor  $a$  (volatilidade) neutro é 1. Se se combina a volatilidade neutra con prestixio asimétrico 4 ou 6, a desaparición dunha das linguas adoita tardar ó redor de 65 unidades de tempo cando se parte dunha repartición xeográfica casual e mesturada dos falantes como na imaxe. Subindo o valor de volatilidade a 2 e mantendo a mesma relación de prestixio, a desaparición tarda entre 180 e 200 unidades de tempo

Aínda así, hai situacións empíricas que parecen confirmar que se comproba a observación de “máis volatilidade, máis estabilidade”, e a segunda resposta á observación provocada polo modelo é que pode realmente ser que haxa situacións nas que o comportamento “volátil” chegue a producir efectos paradoxais. Un exemplo anecdótico: o meu propio comportamento en Galicia adoitaba ser, desde o meu primeiro contacto co galego, pouco volátil e case monolingüe en galego. Nos últimos anos adoptei cada vez máis unha actitude menos consecuenta, máis aberta cara ao cambio, máis “volátil”<sup>18</sup>. Curiosamente, nalgunhas das persoas coas que eu sempre falara asimetricamente en galego (eu en galego, eles en castelán), o cambio de actitude provocou a reacción contraria: desde que lles contesto en castelán, son eles os que cambian ao galego. Isto pode ser moi marxinal, excepcional ou estraño, pero indica algo que é moi importante para a cuestión da modelización do noso comportamento. Non sabemos por que exactamente aqueles falantes reaccionaron así ao meu cambio. Pero podemos afirmar que xeralmente na actualidade o galego pasou de ser unha lingua falada por descoñecemento do castelán a ser unha lingua falada a carón do coñecemento do castelán; é dicir, que é cada vez máis común que falar galego non sexa unha *necesidade* senón unha *escolla consciente* por parte dos falantes, unha opción a carón doutra. Os falantes saben que cando alguén fala galego, a partir dunha certa xeración e en correlación contraria co esquema tradicional dos ámbitos da lingua, faíno porque quere e non porque o necesita. O cambio ao castelán dun galegofalante pódese interpretar, pois, como un xesto dun que sabe dúas linguas cara a outro que só sabe unha —e pode provocar, paradoxalmente, a vontade nese falante de demostrar que el tamén sabe as dúas.

---

<sup>18</sup> As razóns disto son múltiples, en parte son individuais e corresponden ao incremento do contacto con castelanfalantes, en parte tamén son típicos dunha xeración que cos anos ía adoptando actitudes menos monolingües, ao mellor por darse conta do efecto paradoxal que estas podían ter.

Non fai falta que entremos aquí en máis especulacións sobre ese tipo de comportamentos, o que é importante é que nos demostran un feito importantísimo da realidade social: contrariamente á idea antiga de que a comunicación consiste no intercambio de signos máis ou menos comúns entre un emisor e un receptor, na lingüística actual sabemos que os falantes constrúen continuamente hipóteses sobre a cognición dos seus interlocutores, e sabemos que o noso comportamento é altamente estratéxico, incluíndo hipóteses sobre as estratexias dos demais, *emulacións* do outro. Por iso, sabemos que freando menos abruptamente contribuímos a evitar o atasco, e sabemos que cando unha empresa dá boas novas, paradoxalmente, será mellor vender as accións<sup>19</sup>. E podemos saber, por exemplo, que falar unha lingua ou empregar unha forma determinada pode ter efectos variados, pode provocar a imitación, o rexeitamento, ou incluso outras reaccións. Mesturando dúas linguas, o resultado non estará no medio como cando mesturamos dous líquidos de dúas cores: o resultado pode ser de ambas as cores, dunha intermedia ou doutra totalmente diferente.

O comportamento estratéxico, mediato e non inmediato, é un feito que hai que ter en conta á hora de crear modelos de dinámica lingüística, e se un modelo como o de Castelló et al. axuda a provocar ese tipo de reflexións, será considerado como punto de partida e non como punto final na modelización, e indica un camiño digno de ser seguido: a modelización baseada no comportamento de axentes, tendo en conta diferentes constelacións sociais e incluíndo as estratexias dos individuos, tal como se fai por exemplo en modelos baseados na teoría dos xogos.

### 3.3. A base dos datos

Os modelos poden, pois, deixar ver bosques onde os individuos só ven árbores, evolucións de linguas onde os falantes ven actos individuais, e neste sentido poden ser moi útiles. Agora ben, para que a modelización sexa o máis adecuada posible, os modelos necesitan datos fiables e comparables. Entre os modelos presentados, hai casos como o de Minett e Wang, que non traballan con datos concretos e unicamente sinalan a posible aplicación<sup>20</sup>, outros como o de Castelló et alii, que cita casos empíricos concretos de aplicabilidade sen entrar na cuestión concreta da natureza dos datos<sup>21</sup> e outros, como o de Mira e Paredes e o de Abrams e Strogatz, que si traballan con datos concretos. No caso de Mira e Paredes, a cita da autoridade da Real Academia Galega como fonte dos datos sobre a evolución dos galegofalantes entre 1875 e 1975 parece “obxectivar” os datos; no entanto, sabemos que os datos sobre esta evolución anteriores á existencia dunha socioloxía da linguaxe seria son máis ca problemáticos<sup>22</sup>. Se isto é o caso xa nos

<sup>19</sup> O lema “sell on good news, buy on bad news” é coñecido entre os especuladores da bolsa, e ten unha xustificación “real”, con base no mercado principal, no feito de que as empresas polo comportamento cíclico da economía adoitan ter caídas despois do boom e resurreccións despois das caídas. Pero tamén ten unha xustificación no mercado secundario, o das accións, xa que os analistas saben que o coñecemento dese principio vai levar á caída despois de boas noticias e á subida despois de malas noticias: saben iso porque é un principio da bolsa que se autoalimenta. E tamén o saben as empresas, que procuran dar conscientemente noticias esaxeradamente boas en momentos estratéxicos e noticias esaxeradamente malas noutros momentos estratéxicos. O problema do espellismo das estratexias é que poden chegar a destruír as relacións de base: a pura especulación desconecta totalmente da produción real, do mercado principal, e pode chegar a derrubar empresas en realidade sas e con bos produtos.

<sup>20</sup> “Until the predictions of the model have been fitted to empirical data, we hesitate to claim that the probabilities quoted here indicate precisely the likelihood of language maintenance being achieved within an actual community.” (Minett/Wang 2008, 42).

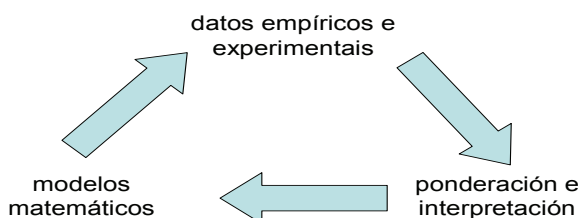
<sup>21</sup> p. ex. “bilingualism in a non-volatile society may have also played a role in the survival of Galician. Indeed, our applet shows that the presence of bilingual agents, in the case of non-volatile systems, contributes to the delay in the extinction of the endangered language, while in the Abrams-Strogatz Model it disappears a lot earlier in time” (Castelló et alii, *no prelo*)

<sup>22</sup> De feito, Mira/Paredes 2005, 1032 citan o libro sobre *Usos lingüísticos* (RAG 1995) coma fonte do seu esquema. Debe tratarse dun erro xa que nese libro se presenta un estudo sincrónico sobre a situación de 1991, e non hai datos sobre a evolución histórica. De feito, non se sabe, pois, de onde veñen os datos do modelo. En todo caso, veñan de onde veñan, só se pode tratar de datos de estimacións xa que non dispoñemos de datos obvios sobre a situación lingüística galega anteriores ao estudo de Ayestarán/De la Cueva de 1974. Desde o *Mapa sociolingüístico de Galicia* (datos de 1991), Galicia é probablemente a comunidade mellor estudada en España na cuestión dos usos lingüísticos; aínda así, como sinalamos nunha recensión sobre o estudo piloto do mapa realizado na comarca ferrolá (Kabatek 1995), incluso ese estudo pormenorizado só repre-

datos referidos a unha única lingua, será aínda moito máis problemático cando se trata de varias linguas, con datos de fontes e natureza moi diversas<sup>23</sup>.

O mellor remedio contra ese problema estará na mellora dos datos; unha mellora que debería resultar xa da colaboración entre lingüistas e modeladores: para unha aplicación fiable dos modelos, o ideal sería partir dunha base empírica ou experimental exclusivamente deseñada para a posible entrada nos modelos. Por exemplo, para non limitarse a meras cifras de falantes e supoñer ou inventar o comportamento dos individuos que explica tal evolución, sería recomendable acadar datos reais ou experimentais que permitisen dar conta do comportamento real dos falantes en situacións concretas, inclusive o comportamento paradoxal do que se falou enriba. Isto permitiría tamén incluír máis factores ou distincións, como a diferenciación interna da lingua ou os cambios na relación de prestixio entre as linguas ao longo do tempo. Pero isto parece máis doado do que é, e non hai que enganarse: o camiño desde o comportamento individual á evolución da sociedade é moi difícil de trazar, e non é casual que tal cousa practicamente nunca se lograra no pasado e que por iso haxa recursos a metáforas coma a da “man invisible” que máis ca explicar, encobren ese camiño. Ou hai modelos que calculan e visualizan só o que xa é obvio. A verdadeira tarefa dunha modelización adecuada presupón a creación de escenarios microlingüísticos que deixen ver a *calidade* dos fenómenos de interacción e unha análise da representatividade deses fenómenos.

Con esa base, a “espiral heurística” da que falamos sería ampliada, non se limitaría á interacción entre os modeladores e os lingüistas, dando os modeladores o modelo e os lingüistas as ideas dos posibles factores determinantes. Habería que engadir un terceiro elemento, os datos empíricos e a súa ponderación e interpretación, como se representa na fig. 6.



**Fig. 6.** Dos datos empíricos á modelización pasando pola ponderación e interpretación dos datos

Estamos aínda lonxe da aplicación deste triángulo a situacións complexas do contacto de linguas, pero creo que iluminar a vía entre datos empíricos reais e modelos é un dos verdadeiros obxectivos dunha renovada lingüística do século XXI. Todo o que só aparentemente establece ese vínculo sen datos reais é máis suposición ca outra cousa, e nalgúns casos parece ata ser un engano.

Ao mellor, para non enganarnos, non habería que comezar coa idea de termos que ser capaces de modelar comunidades enteiras e contactos entre linguas: habería que contentarse quizais con situacións máis reducidas, pequenos grupos, familias, etc. O soño da socioloxía de establecer unha ponte real entre “micro” e “macro” pode chegar a ser realidade no século XXI, pero unicamente se confesamos cunha certa humildade que os pasos por dar aínda quedan pendentes.

senta datos sobre o que os falantes *din* nunha situación concreta de enquisa, e sabemos que entre *belief* e *behavior*, o que se cre e se di e o que se fai, pode haber abismos.

<sup>23</sup> No breve artigo de Abrams e Strogatz (2003), díse que se traballou con datos de 42 linguas sen dicir cales nin entrar na cuestión da calidade e comparabilidade dos datos. Os únicos catro casos presentados no artigo son os que reproducimos na fig. 2.



## 4. CONCLUSIONES

Nas reflexións que presentamos neste traballo partimos dun certo entusiasmo que marca a lingüística actual: as novas metodoloxías cuantitativas, no caso aquí presentado, lévannos a crer que por fin os lingüistas somos capaces de resolver unha serie de problemas abertos que sempre nos preocuparon. A ponte entre comportamento individual e evolución dunha lingua, a miúdo presuposta, pero raras veces feita visible, parece ao noso alcance grazas ás novas técnicas de modelización e de cálculo. Pero a todo entusiasmo deste tipo segue, na sobriedade, a reflexión sobre os avances reais, e alén dos novos logros, hai que sinalar tamén os límites. Identificamos basicamente tres ámbitos de tales límites que afectan en diferentes graos os diferentes modelos que describen a dinámica das linguas:

- o problema dos falsos supostos
- o problema da simplificación
- o problema dos datos

Os tres problemas están relacionados, pero ata certo punto son separables. A solución dos tres reside no último, na base de datos empíricos adecuada para o estudo. Unha vez que dispoñamos de datos auténticos que deixen ver os fenómenos reais da interacción lingüística, poderán entrar nun modelo; levarán a modelos máis axeitados, non baseados na intuición que teñen os modeladores sobre unha situación dinámica, senón en comportamentos reais. Eses modelos terán que dar conta dos factores concretos da interacción, e prescindir de simplificacións inadecuadas; así, por exemplo, no canto de supoñer que todos teñen o mesmo contacto con todos, haberá que ter en conta redes de contacto como as empiricamente comprobadas; no canto de supoñer que hai territorios homoxéneos, haberá que dar conta das variacións diatópicas, distráticas e diafásicas; no canto de pensar que a relación de prestixio entre as linguas é estable ao longo do tempo, haberá que crear modelos dinámicos nos que a relación de atractividade respectiva das linguas de contacto sexa variable<sup>24</sup>. Finalmente, o terceiro problema, o dos falsos supostos, tamén se resolverá cando o vínculo cos datos empíricos non sexa suposto ou inventado, senón real. E mesmo tendo en conta todo isto, o cálculo matemático evidentemente non nos resolverá todas as cuestións abertas da descrición do contacto e do desprazamento lingüísticos, pero axudará a comprender a distinguir os factores máis relevantes. Neste sentido, estamos a vivir tempos apaixonantes e agardamos con ansia as novas visións que nos traerá a lingüística deste século.

## REFERENCIAS BIBLIOGRÁFICAS

- Abrams, Daniel M. / Steven H. Strogatz (2003): "Modeling the dynamics of language death", *Nature* 424, 900.
- Ayestarán Aranaz, Margarita / Justo de la Cueva Alonso (1974): *Las familias de la provincia de Pontevedra (Galleguidad y conflicto lingüístico)*. Sevilla: Instituto de Ciencias de la Familia.
- Axelrod, Robert (2006): "Agent-based modeling as a bridge between disciplines", en Leigh Tesfatsion / Kenneth L. Judd (eds.), *Handbook of Computational Economics Vol. 2: Agent-Based Computational Economics*. Amsterdam: North Holland/Elsevier, 1565–1584.

<sup>24</sup> Neste contexto, queremos sinalar unha certa inxenuidade coa que nalgúns casos dos traballos presentados se fala na intervención para cambiar unha situación: se no traballo de Minett e Wang (2008: 28) se di, citando a Crystal, que incrementar o prestixio aumenta as posibilidades de supervivencia dunha lingua, é como se digo que unha persoa sa vai enfermarse menos que unha persoa doente. Se se di, ademais, que o que aumenta as posibilidades de supervivencia é tamén incrementar o número de monolingües, hai que avisar do problema ético á hora de pensar na creación artificial de "guetos lingüísticos". Como di Mufwene (2004: 219), citado por Minett e Wang: "Linguists concerned with rights of languages must ask themselves whether these rights prevail over the rights of the speakers to adapt competitively to their new socioeconomic ecologies."



- Blythe, Richard / William Croft (2010): "Can a science-humanities collaboration be successful?", *Adaptive Behavior* 18, 12–20.
- Bornkessel-Schlesewsky, Ina / Matthias Schlesewsky (2009): *Processing Syntax and Morphology: A Neurocognitive Perspective*. Oxford: Oxford University Press.
- Castelló, Xavier (2010): *Collective phenomena in social dynamics: consensus problems, ordering dynamics and language competition*, Tese de doutoramento UIB <http://ifisc.uib-csic.es/publications/download.php?fid=4001>.
- Castelló, Xavier et alii (2008): "Modelling language competition: Bilingualism and complex social networks", en Andrew D.M. Smith / Kenny Smith / Ramón Ferrer i Cancho (eds.), *The evolution of language. Proceedings of the 7th International Conference (EVOLANG7)*. Singapore: World Scientific Publishing Co., 59–66.
- Castelló, Xavier / Lucía Loureiro-Porto / Maxi San Miguel (no prelo): "Agent-based models of language competition", en *International Journal of the Sociology of Language*, Special issue on *Models and language contact*.
- Crystal, David (2000): *Language Death*. Cambridge: CUP.
- Fishman, Joshua A. (1991): *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon, Philadelphia: Multilingual Matters.
- Fishman, Joshua A. (ed.) (2001): *Can Threatened Languages Be Saved? Reversing Language Shift Revisited: A 21st Century Perspective*. Clevedon, Buffalo: Multilingual Matters.
- Kabatek, Johannes (1995): Recensión de: Seminario de Sociolingüística da Real Academia Galega (ed.): *Estudio sociolingüístico da comarca ferrolá. Fase previa ó Mapa Sociolingüístico de Galicia*. A Coruña: Real Academia Galega 1993 (*Cadernos de Lingua*, Anexo 1), *Sociolingüística* 9 (1995), 160–162.
- Kabatek, Johannes (2000): *Os falantes como lingüistas*. Vigo: Xerais.
- Kabatek, Johannes (2005): "Über Trampelpfade, sichtbare Hände und Sprachwandelprozesse", en Thomas Stehl (ed.), *Unsichtbare Hand und Sprecherwahl. Typologie und Prozesse des Sprachwandels in der Romania*. Tübingen: Narr, 155–174.
- Kabatek, Johannes / Lucía Loureiro-Porto (no prelo): "Mathematic models meet linguistic data and vice versa", *International Journal of the Sociology of Language*, Special issue on *Models and language contact*.
- Katsumori, Makoto (2010): "Heisenberg on Science, Language, and the Question of Objectivity", *Energeia* 2, 1–7.
- Keller, Rudi (1990/1994): *Sprachwandel. Von der unsichtbaren Hand in der Sprache*. Tübingen: Francke.
- Minett, James W. / William S.-Y. Wang (2008): "Modelling endangered languages: The effects of bilingualism and social structure", *Lingua* 118, 19–45.
- Mira, Jorge B. / Ángel Paredes (2005) "Interlinguistic similarity and language death dynamics", *Europhysics Letters* 69 (6), 1031–1034.
- Mira, Jorge / Luís F. Seoane / Juan J. Nieto (2011): "The importance of interlinguistic similarity and stable bilingualism when two languages compete", *New Journal of Physics* 13, [http://iopscience.iop.org/1367-2630/13/3/033007/pdf/1367-2630\\_13\\_3\\_033007.pdf](http://iopscience.iop.org/1367-2630/13/3/033007/pdf/1367-2630_13_3_033007.pdf).
- Mufwene, Salikoko (2004): "Language birth and death", *Annual Review of Anthropology* 33, 201–222.
- Real Academia Galega. Seminario de Sociolingüística (1995): *Usos lingüísticos en Galicia*. A Coruña: RAG.
- Trudgill, Peter (1972): "Sex, covert prestige and linguistic change in the urban British English of Norwich", *Language in Society* 17, 61–88.